Impact of AI on Autonomous Driving

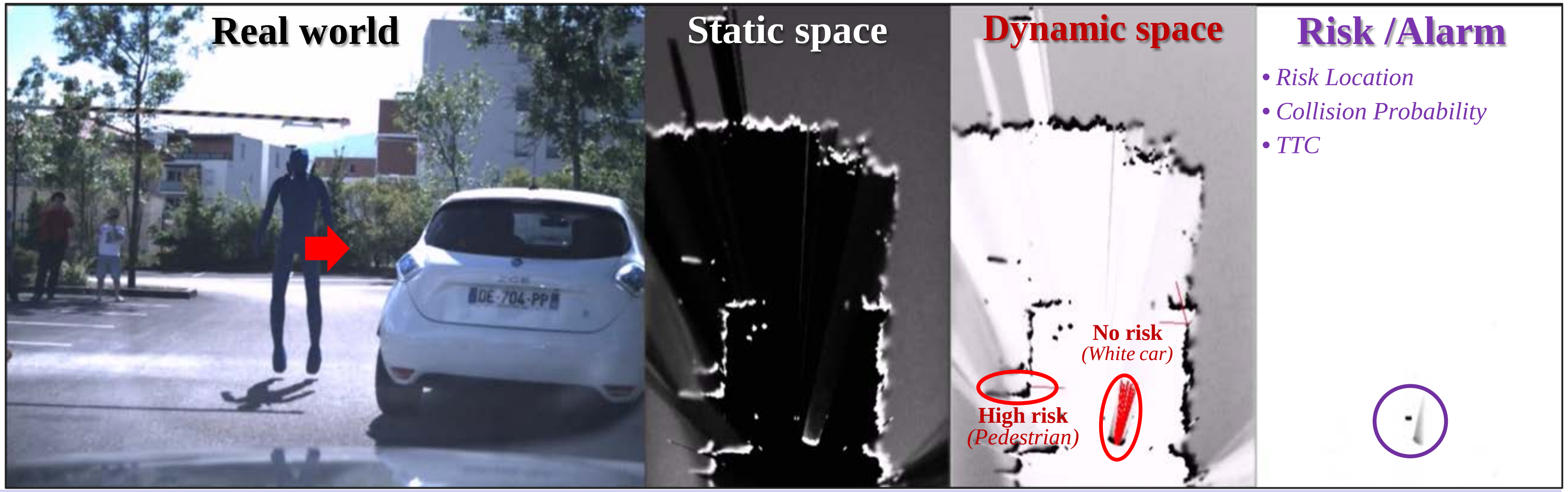
Christian Laugier

Impact of AI on Autonomous Driving

Christian LAUGIER, PhD & Dr Es Science
Research Director at Inria & Scientific Advisor for Probayes and for Baidu China
Inria Chroma team & IRT Nanoelec
christian.laugier@inria.fr

Contributions from
L. Rummelhard, A. Negre, N. Turro, J.A. David, J. Lussereau, T. Genevois, C. Tay Meng Keat, S. Lefevre, O. Erkent, D. Sierra-Gonzalez

ADAS & Autonomous Driving



Keynote talk @ WRC 2019 Conference
Beijing Etrong International Exhibition & Convention Center, Beijing, China, August 21-23 2019

informatics mathematics

NANOELEC

Technology status & Ongoing challenges for AVs

- **Strong involvement of Car Industry & GAFA + Large media coverage + Increasing Governments supports**
- **An expected market of 515 B€ at horizon 2035** *(~17% world automobile market, Consulting agency AT Kearney, Dec 2017)*
- **But Legal & Regulation issues are still unclear … idem for Technologies Validation & Certification issues !**

=> Numerous experiments in real traffic conditions since 2010 (Disengagement reports ⇔ Insights on system maturity)

=> But still insufficient … Realistic Simulation & Formal methods are also under development (e.g. EU Enable-S3)

Tesla Autopilot L2 with Radar & Mobileye/Intel
Commercial ADAS product => ***Tested by customers !!***

Numerous EU projects in last decades
Cybus, 3 months experiment, low speed
La Rochelle 2012 (France)

Dense 3D mapping & Numerous vehicles
10 years R&D, 8 millions km covered since 2010 & 25 000 km/day

Drive Me trials (Volvo, 2017)
- *100 Test Vehicles in Göteborg, 80 km, 70km/h*
- *No pedestrians & Separations between lanes*

"Self-Driving Taxi Service L3" testing in US *(Uber, Waymo)* & Singapore *(nuTonomy)*

⇒ ***Autonomous Mobility Service****, Numerous Sensors + "Safety driver" during testing (take over in case*
⇒ ***Uber:*** *System testing since 2017, Disengagement every 0.7 miles in 2017 (improved now)*
⇒ ***Waymo:*** *1st US Self Driving Taxi Service launched in Phoenix in Dec 2018*
⇒ ***Disengagement reports provide insights on the technology maturity***

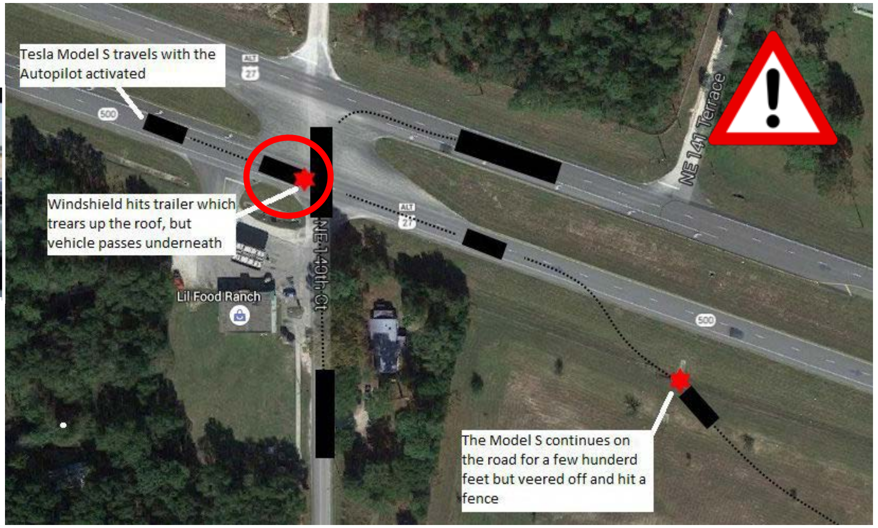
Millions of miles driven since 2010 (Google, Tesla, Waymo, Uber, ..)

Several benign & serious accidents in past few years

Safety is still not guaranteed !

Fatal accidents involving AVs – *Perception failure*

❑ Tesla driver killed in a crash with Autopilot "level 2" active *(ADAS mode) – May 2016*



- ✓ ***The Autopilot <u>failed to detect</u> a white moving truck, with a brightly lit sky*** *(Camera Mobileye + Radar)*
- ✓ ***The human driver was not vigilant***

❑ Self-driving Uber L3 vehicle killed a woman in the first fatal crash involving a pedestrian

Temple, Arizona, March 2018

- ✓ ***Despite the presence of multiple sensors, the <u>perception system failed to detect the pedestrian & didn't disengage</u>***
- ✓ ***The Safety Driver reacted too lately*** *(1s before the crash)*

AVs have to face two main challenges

1. The need for **Robust**, **Self-diagnosing** and **Explainable** ***Embedded Perception***

Video source: AutoPilot Review @ youtube.com

Video scenario: *The Tesla perception system (cameras + radar) failed to detect the barrier on the left road side. The vehicle was steering towards the barrier when the* ***human driver took over*** *for avoiding the collision.*

AVs have to face two main challenges

1. The need for **Robust**, **Self-diagnosing** and **Explainable** ***Embedded Perception***

Video source: AutoPilot Review @ youtube.com

Video scenario: *The Tesla perception system (cameras + radar) failed to detect the barrier on the left road side. The vehicle was steering towards the barrier when the* ***human driver took over*** *for avoiding the collision.*

2. The need for **Understandable** ***Driving Decisions*** (sharing the road with human drivers)

Human drivers actions are determined by a complex set of interdependent factors difficult to model (e.g. intentions, perception, emotions …)

$\Rightarrow$ ***Any human driver behavior prediction is inherently uncertain***

$\Rightarrow$ ***AV have to reason about uncertain intentions of surrounding vehicles***

Video source: The Telegraph

Video scenario (from dash cam of a bus moving behind of the Waymo AV):

- *The* ***Waymo AV*** *is blocked by an obstacle and it decides to change lane*
- *The* ***bus driver*** *misunderstood the Tesla's intention and didn't yield*
- *The two vehicles collided*

Perception & Decision-making requirements for AVs

Dynamic Scenes Understanding & Navigation Decisions

Situation Awareness & Decision-making

⇒ *Sensing + Prior knowledge + Interpretation*

⇒ *Navigation strategy selection (planning & control)*

Embedded Perception & Decision-making for Safe Intentional Navigation

Dealing with unexpected events

Anticipation & Risk Prediction

for avoiding an upcoming collision with "something"

=> High reactivity & Reflexive actions

=> Focus of Attention & Sensing

=> Collision Risk estimation + Avoidance strategy

Main features

✓ Dynamic & Open Environments => *Real-time processing & Reactivity (several reasoning levels are required)*

✓ Incompleteness & Uncertainty => *Appropriate Model & Algorithms (probabilistic approaches)*

✓ Sensors limitations (no sensor is perfect) => *Multi-Sensors Fusion*

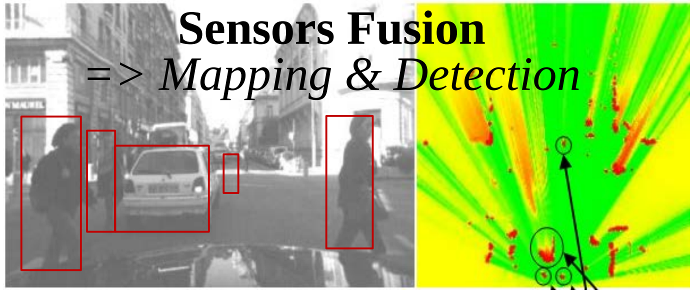
✓ Hardware / Software integration => *Satisfying Embedded constraints*

✓ Human in the loop (mixed traffic) => ***Human Aware Decision-making process (AI based technologies)***

Taking into account Interactions + Behaviors + Social rules (including traffic rules)

Paradigm 1: Embedded Bayesian Perception

Embedded Multi-Sensors Perception
⇒ Continuous monitoring of the dynamic environment



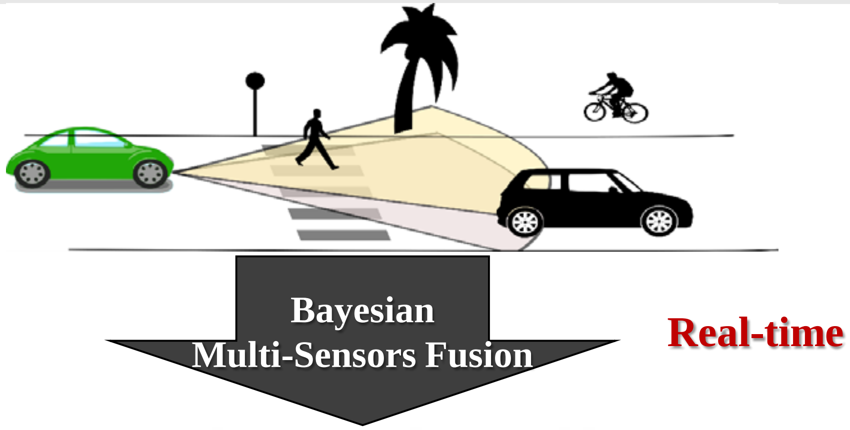
❑ Main challenges

- ✓ *Noisy data, Incompleteness, Dynamicity, Discrete measurements*
- ✓ ***Strong Embedded & Real time constraints***

❑ Our Approach: Embedded Bayesian Perception

- ✓ *Reasoning about **Uncertainty** & **Time window** (Past & Future events)*
- ✓ ***Improving robustness** using **Bayesian Sensors Fusion***
- ✓ ***Interpreting the dynamic scene** using **Contextual & Semantic** information*
- ✓ ***Software & Hardware integration** using GPU, Multicore, Microcontrollers…*

Bayesian Perception : Basic idea



- **Multi-Sensors Observations**

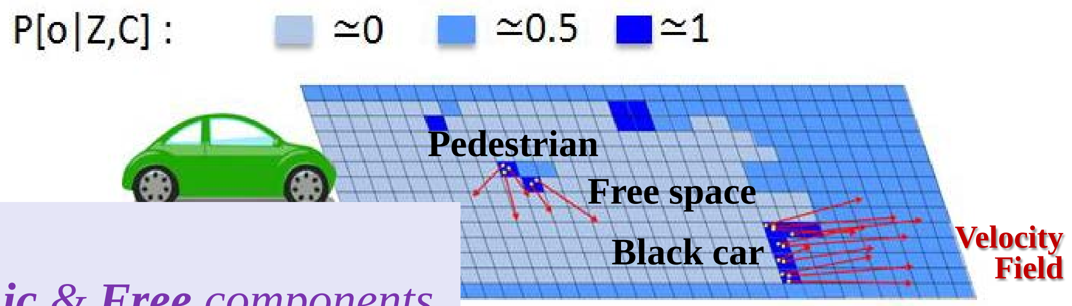
  Lidar, Radar, Stereo camera, IMU …

- **Probabilistic Environment Model**

  ✓*Probabilistic Sensor Fusion*

  ✓*Probabilistic representation of Occupancies & Velocities*

  ✓*Prediction models*



Concept of "Dynamic Probabilistic Grid"

⇒ *Clear distinction between **Static** & **Dynamic** & **Free** components*

⇒ *Occupancy & Velocity probabilities*

⇒ *Designed for Highly Parallel Processing*

⇒ *Embedded models for **Motion Prediction***

- **Main philosophy**

  Reasoning at the grid level as far as possible for both :

  - ***Improving Efficiency & Reactivity to unexpected events*** *=> Highly parallel processing & high frequency*
  - ***Avoiding most of traditional object level processing problems*** *(e.g. detection errors, wrong data association…)*

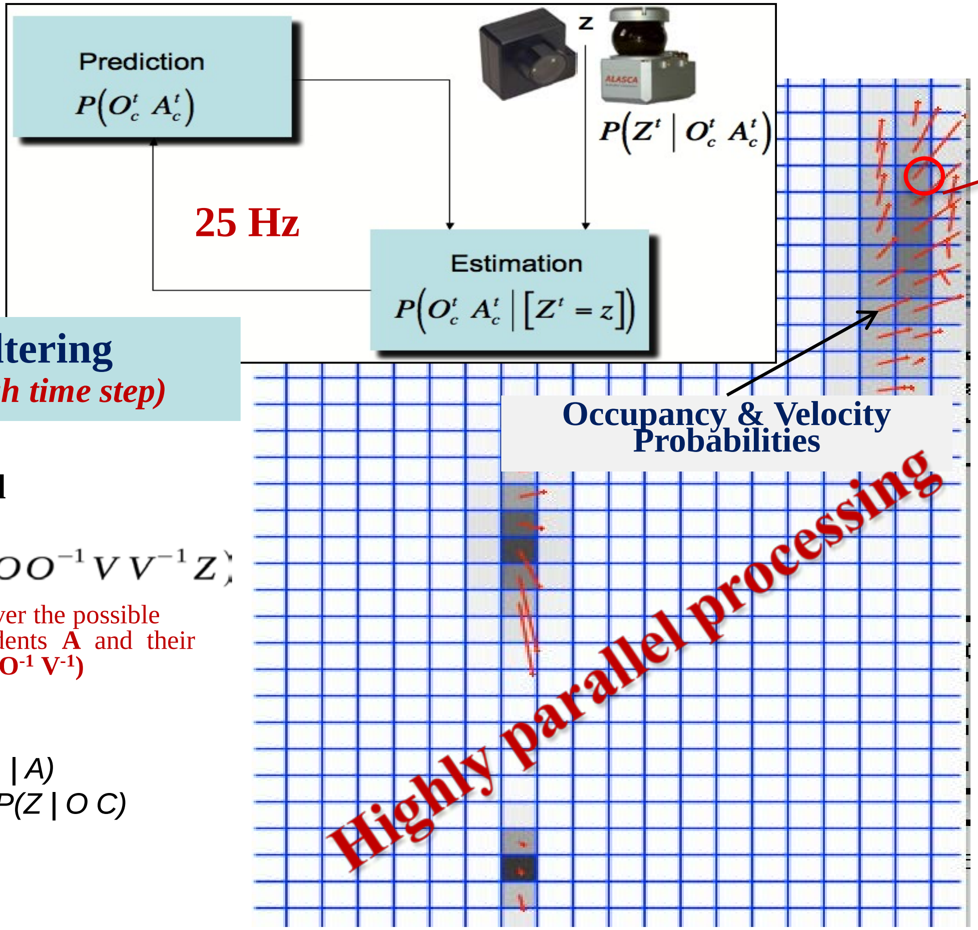
Dynamic Probabilistic Grid – *Implementation outline*

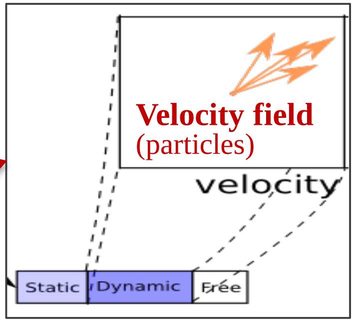
⇒ A more and more popular approach for Autonomous Vehicles
⇒ A clear distinction between Static & Dynamic & Free components

Pioneer concept of "Bayesian Occupancy Filter" (Inria)

[PhD Thesis Coué 2005] [Coué & Laugier IJRR 05] [Laugier et al ITSM 2011] [Laugier, Vasquez, Martinelli Mooc uTOP 2015]

Sensor Fusion





Bayesian Filtering
(Grid update at each time step)

Solving for each cell

$$P(O\,V \mid Z\,C) = \lambda \sum_{A\,O^{-1}\,V^{-1}} P(C\,A\,O\,O^{-1}\,V\,V^{-1}\,Z)$$

Sum over the possible antecedents **A** and their states **(O^{-1} V^{-1})**

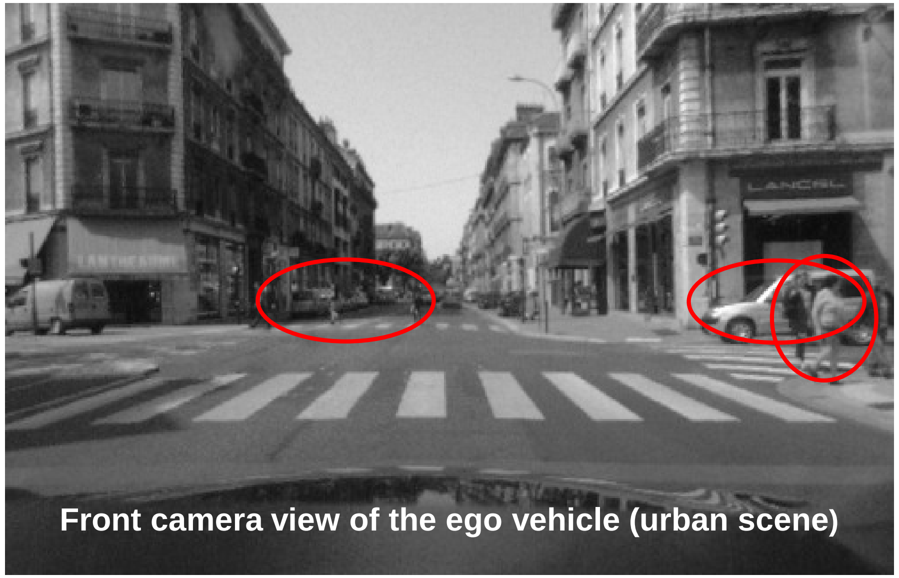
Joint Probability decomposition:

$P(C\ A\ O\ O^{-1}\ V\ V^{-1}\ Z) = P(A)\ \ P(O^{-1}\ V^{-1} \mid A)$
$P(O\ V \mid O^{-1}\ V^{-1})\ \ P(C \mid A\ V)\ \ P(Z \mid O\ C)$

- ***Bayesian Occupancy Filter Approach***
  - *=> Patented by Inria & Probayes (2005)*
  - *=> Commercialized by Probayes*
  - *=> Robust to sensing errors & occultation*
  - *=> Designed for highly parallel processing*
- ***Real-time embedded implementation***
  - *=> Nvidia GPU, 2014*
- ***Used by:*** *Toyota, Denso, Probayes, Renault, EasyMile, IRT Nanoelec / CEA*
- ***Industrial licenses 2018:*** *Toyota + XX (confidential)*

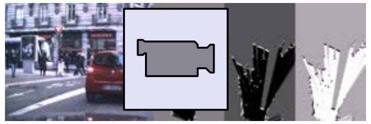
Dynamic Probabilistic Grid – *Main Features*

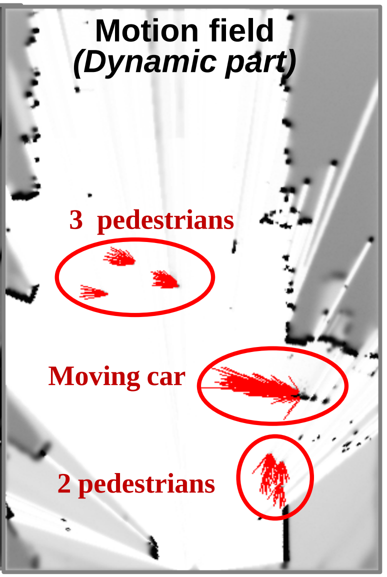
=> Exploiting the dynamic information for a better understanding of the scene



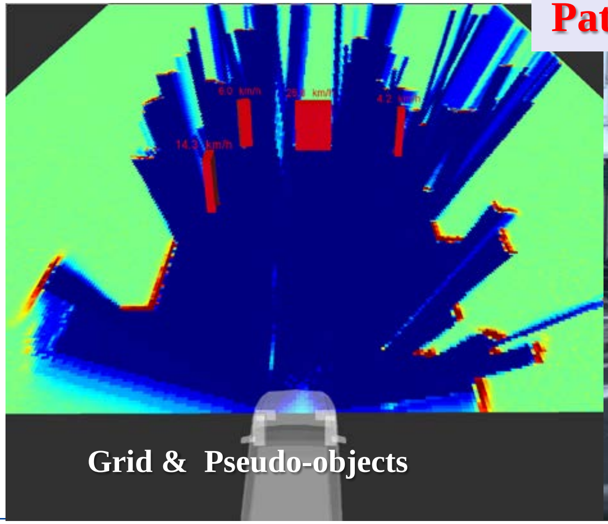
Sensors data fusion
+
Bayesian Filtering
+
Extracted Motion Fields

HSBOF
1st Embedded & Optimized version (patent 2014)

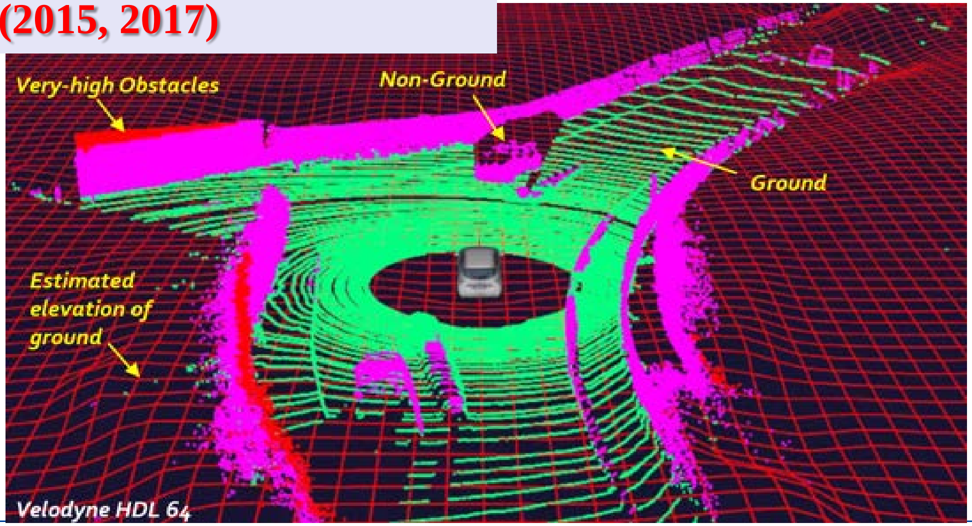




Patented Improvements & Implementations (2015, 2017)



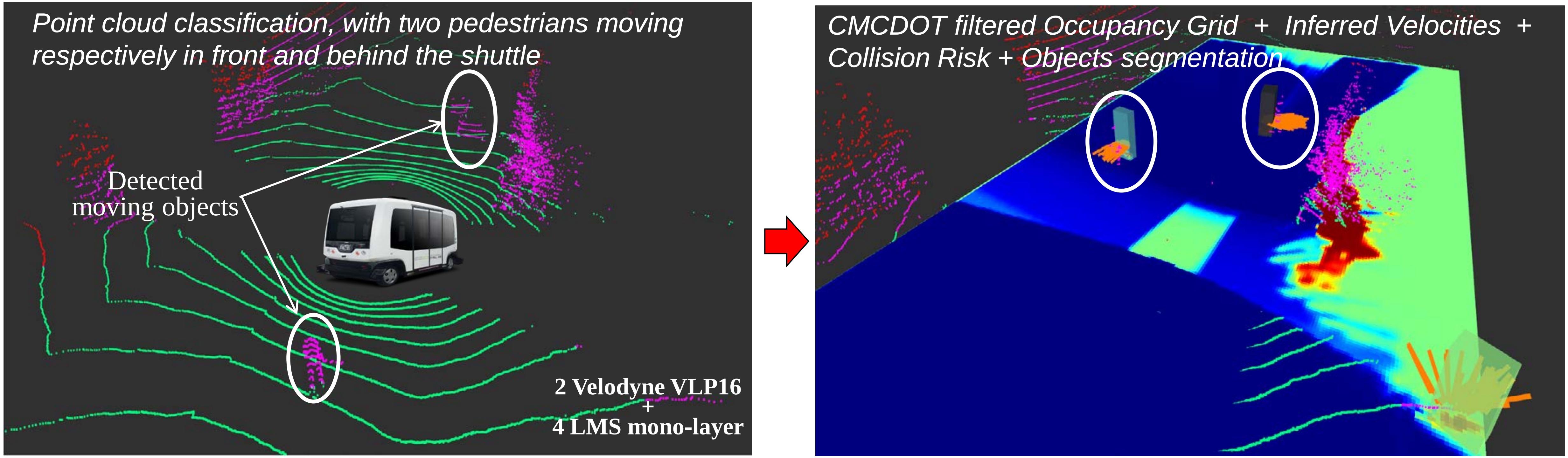
Detection & Tracking + Moving Objects Classification
=> CMCDOT 2015 (including a "Dense Occupancy Tracker")



Ground Estimation & Point Cloud Classification
(patent 2017)

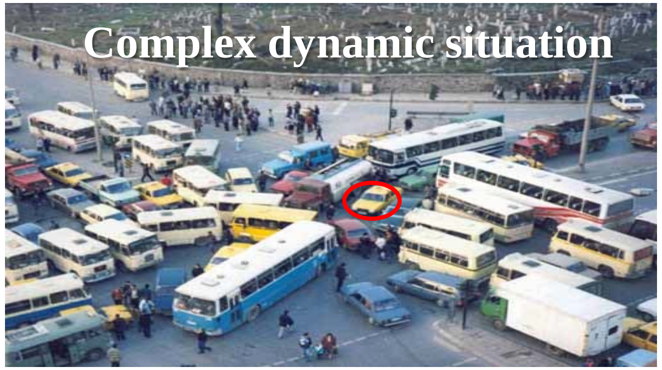
Towards System Integration on a commercial vehicle

- **POC 2018: Complete system implemented on Nvidia TX1**, and easily connected to the shuttle system network *in a few days* (using ROS)
- **Shuttle sensors data** has been fused and processed in **real-time**, with a successful Detection & Characterization of the **Moving & Static Obstacles**
- **Full integration on a commercial product** under development with an industrial company (confidential)



Paradigm 2: Collision Risk Assessment & Decision-making

=> Decision-making for avoiding Pending & Future Collisions



- **Main challenges**

  Uncertainty, Partial Knowledge, World changes, Real time

  Human in the loop + Unexpected events + Navigation Decision based on Perception & Prior Knowledge

- **Approach: *Prediction + Risk Assessment + Bayesian Decision-making***
  - ✓ *Reason about **Uncertainty & Contextual Knowledge** (using **History** & **Prediction**)*
  - ✓ *Estimate **Probabilistic Collision Risk** at a given **time horizon t+δ** (δ = a few seconds)*
  - ✓ *Make **Driving Decisions** by taking into account the **Predicted behavior** of all the observed surrounding traffic participants (cars, cycles, pedestrians …) & **Social / Traffic rules***

- **Decision-making: *Two types of "collision risk" have to be considered***
  - ✓ ***Short-term collision risk** => Imminent collisions with "something" (unclassified), time horizon <3s, conservative hypotheses*
  - ✓ ***Long-term collision risk** => Future potential collisions, horizon >3s, Context + Semantics, Behavior models*

Concept 1: Short-term collision risk – *Basic idea*

=> How to deal with unexpected & unclassified events (i.e. "something" is moving ahead) ?
=> Exploit previous observations for anticipating future objects motions & related potential future collision

Autonomous Vehicle (Cycab)

Parked Vehicle (occultation)

Pioneer Results (2005)

[PhD Thesis C. Coué 2004]
[Coué & Laugier IJRR 05]

Thanks to the prediction capability of the BOF technology, the Autonomous Vehicle "anticipates" the pedestrian motion and brakes *(even if the pedestrian is temporarily hidden by the parked vehicle)*

Short-term collision risk – *Main features*

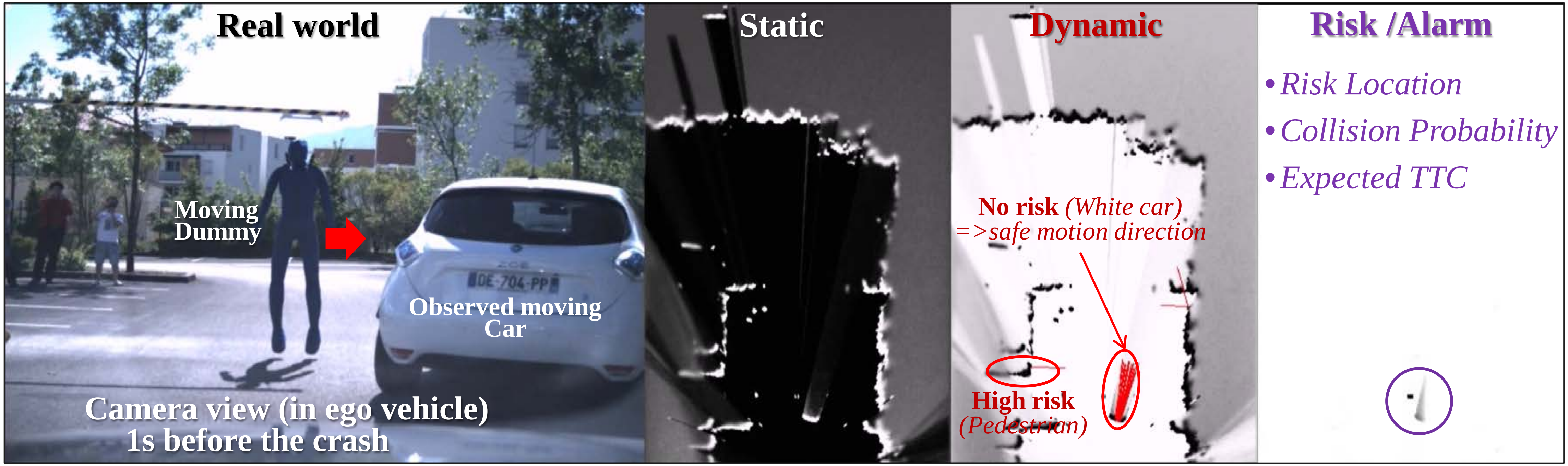
=> Grid level & Conservative motion hypotheses (proximity perception)

Proximity perception: d <100m and t <5s
δ= 0.5 s => Precrash
δ= 1 s => Collision mitigation
δ > 1.5s => Warning / Emergency Braking

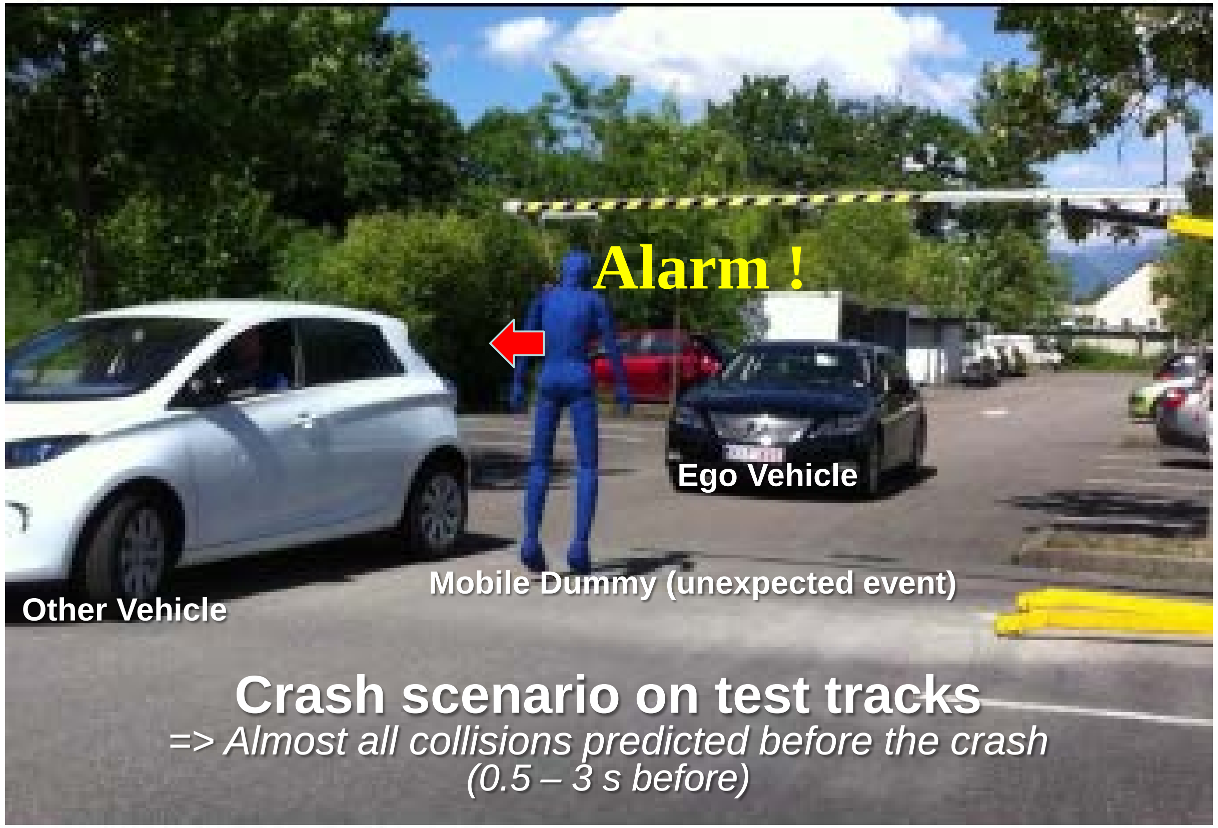
Main Features

- *Detect **"Upcoming potential Collisions"** a few seconds ahead (3-5s) in the Dynamic Grid*
- *Risky situations are **both localized in <u>Space & Time</u>** (under conservative motion hypotheses)*
- *Resulting information is used for choosing **Avoidance Maneuvers***

System output (real time)



Detect potential upcoming collisions & Reduce drastically false alarms



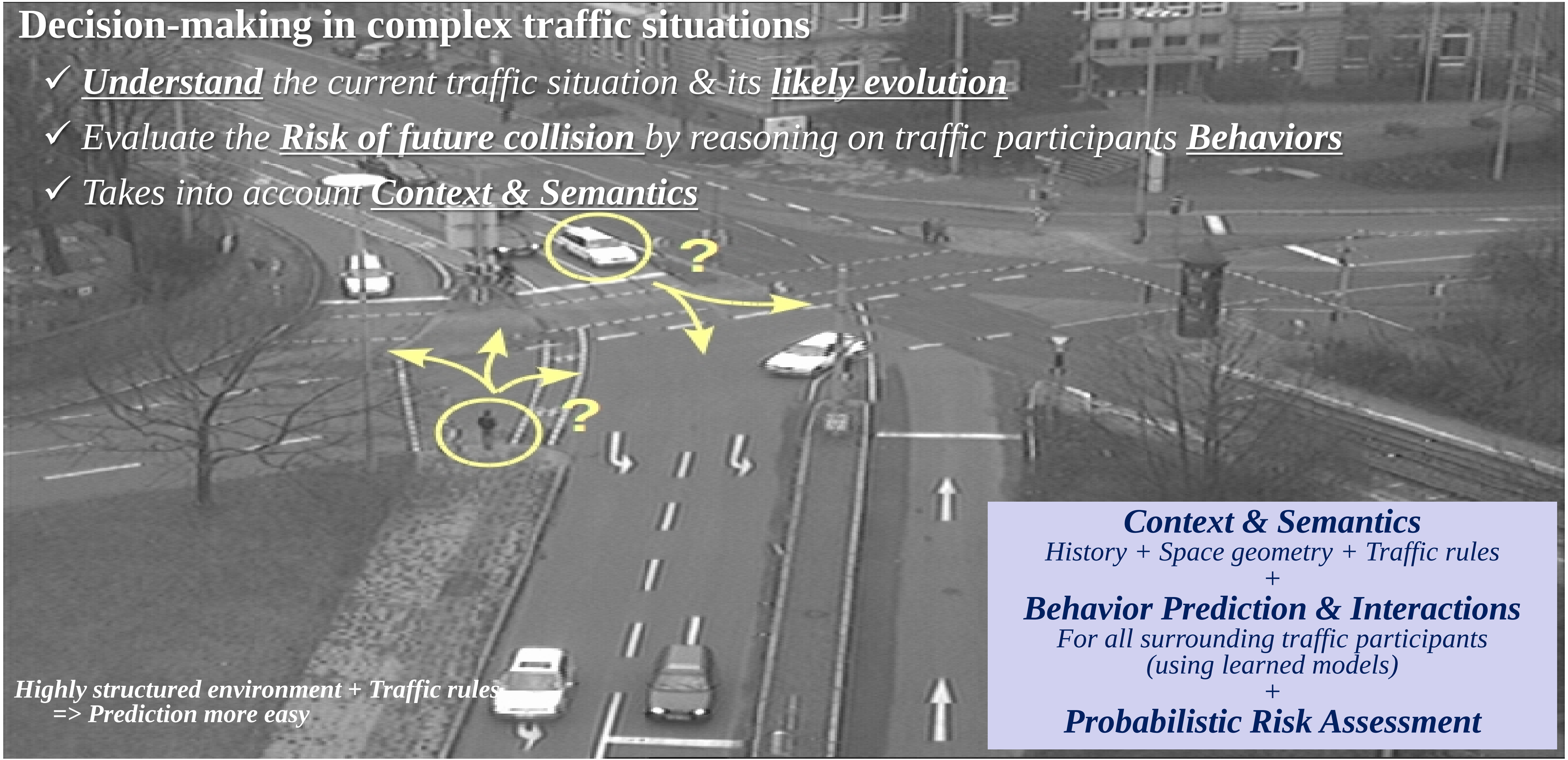
Collision Risk Assessment (video – 0:45)

- **Yellow** => *time to collision: 3s*
- **Orange** => *time to collision: 2s*
- **Red** => *time to collision: 1s*

Concept 2: **Long-term Collision Risk** *(Object level)*

*=> Increasing time horizon & complexity using **Context & Semantics***

*=> Key concepts: **Behaviors Modeling & Prediction + Traffic Participants Interactions***



Decision-making in complex traffic situations

- ✓ ***Understand*** *the current traffic situation & its* ***likely evolution***
- ✓ *Evaluate the* ***Risk of future collision*** *by reasoning on traffic participants* ***Behaviors***
- ✓ *Takes into account* ***Context & Semantics***

Highly structured environment + Traffic rules
=> Prediction more easy

Context & Semantics
History + Space geometry + Traffic rules
+
Behavior Prediction & Interactions
For all surrounding traffic participants
(using learned models)
+
Probabilistic Risk Assessment

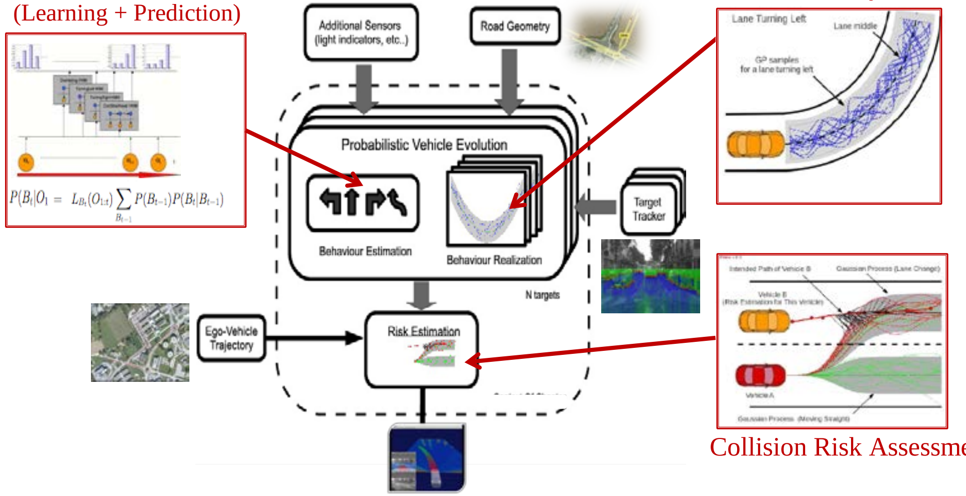
Behavior-based Collision risk *(Object level)*

=> Increased time horizon & complexity + Reasoning on Behaviors & Interactions

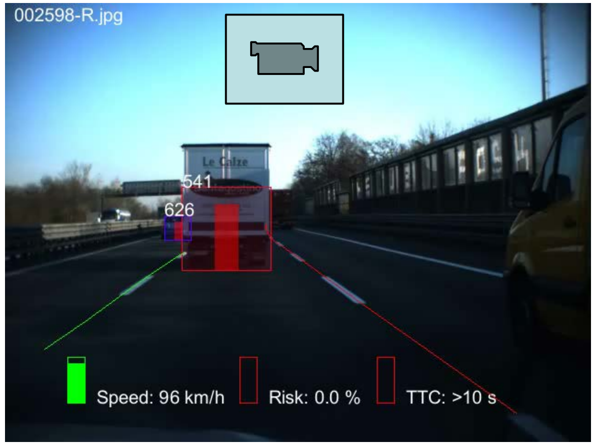
❑ Trajectory prediction & Collision Risk *=> Patent 2010 (Inria, Toyota, Probayes)*

Behaviors
(Learning + Prediction)

Behaviors to Trajectories



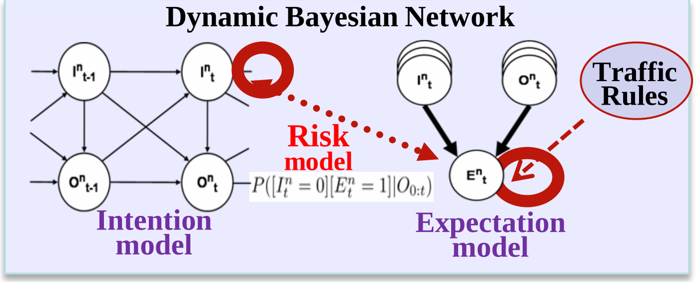
Collision Risk Assessment



Courtesy
Probayes

Cooperation still on-going (R&D contracts + PhD)

❑ Intention & Expectation *(Mixed Traffic & Interactions) => Patents 2012 (Inria - Renault) & 2013 (Inria - Berkeley)*



Human-like reasoning

Cooperation still on-going (R&D contracts + PhD)

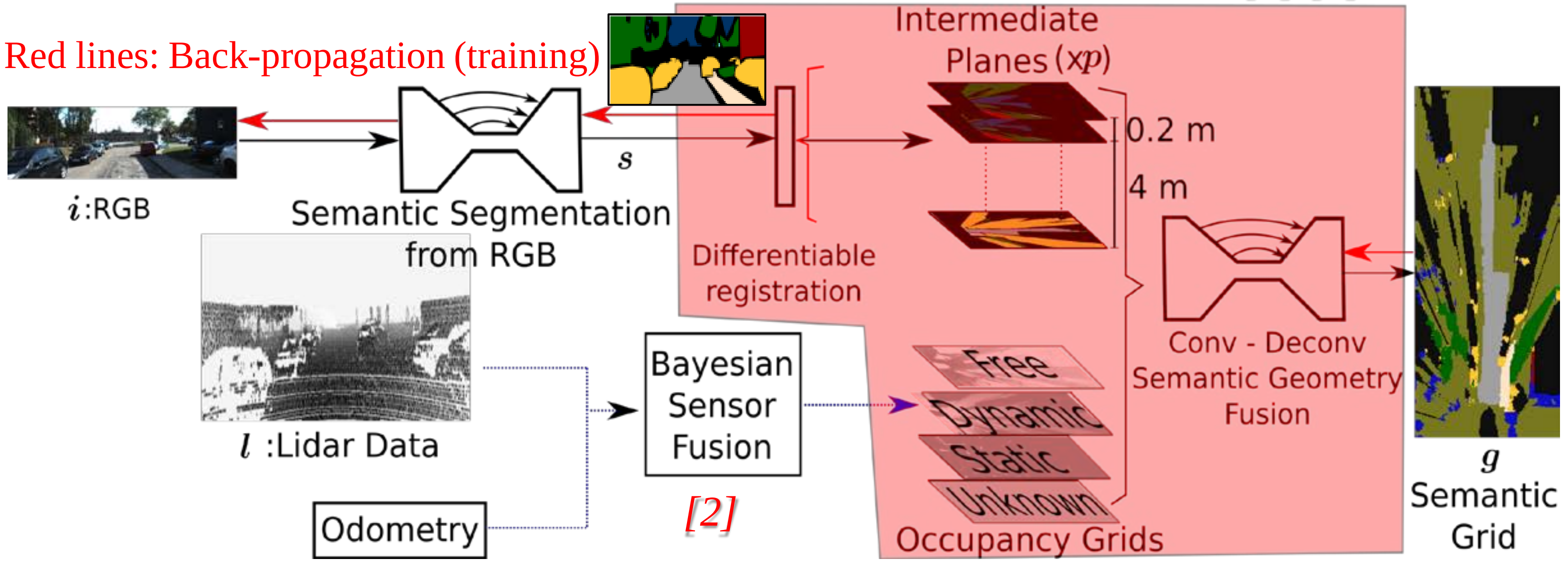
Paradigm 3: Models improvements using Machine Learning
Perception level – Concept of "Semantic Grid"

Objective: *Construct a Bird's Eye View representation of the environment around the vehicle, with **Semantics** (cars, pedestrians, roads, buildings …) recorded in each cell of the **Dynamic Occupancy Grid***

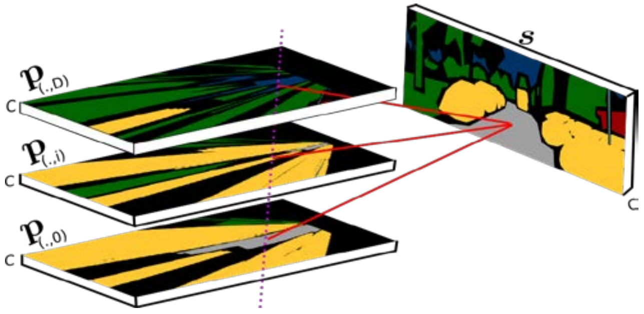
Approach: *A new **Hybrid Sensor Fusion approach** combining **Bayesian Perception & Deep Learning***
[1] [2] + Patent 2019 (Inria, Toyota)

Semantic Grid Network *[1] [3]*

Red lines: Back-propagation (training)



Implementation: Segnet Cuda-GPU + Kitti dataset

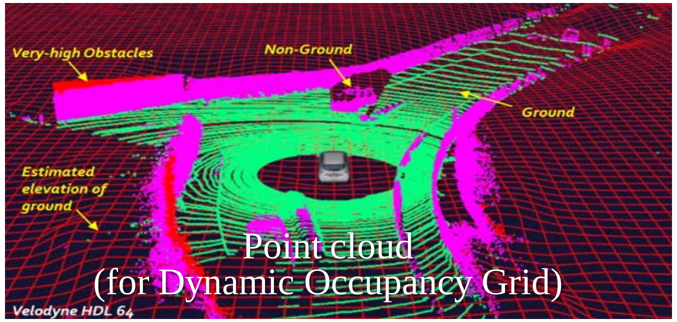


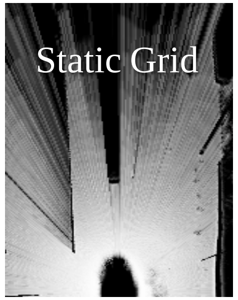
Intermediate layers (~20 layers)

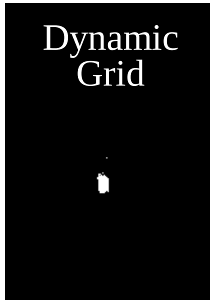
- Intuition: Learns the approximate heights of the classes/objects
- No 3D reconstruction required
- Known "Camera / 3D points" calibration
- Less sensitivity to calibration errors

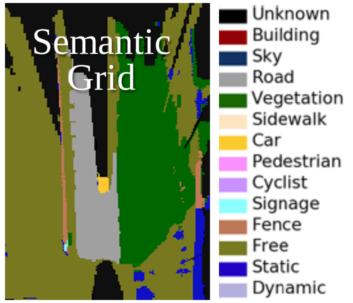
[1] Semantic grid estimation with a Hybrid Bayesian and Deep Neural Network approach, 0. Erkent et al., IEEE IROS 2018
[2] Conditional Monte Carlo Dense Occupancy Tracker, Rummelhard et al., ITSC 2015
[3] Segnet: A deep convolutional encoder-decoder architecture for image segmentation, Badrinarayanan et al., IEEE PAMI 39(12) 2017

Semantic Grids: *Main Characteristics & Properties*









- **Bayesian Sensor Fusion**
  - Real-time implementation available + Accurate and Repeatable results (important for Industry) *[2]*
  - Dense output *independent* of input sensor type … *but* *limited classification capacities*

- **Deep Learning**
  - Good at classification (semantic segmentation)
  - Fast implementation (~real-time) in Cuda available *[3]*
  - Trainable for new situations with new data … *but strongly depends on input data type (e.g. RGB images, point clouds)*

[1] Semantic grid estimation with a Hybrid Bayesian and Deep Neural Network approach, 0. Erkent et al., IEEE IROS 2018
[2] Conditional Monte Carlo Dense Occupancy Tracker, Rummelhard et al., ITSC 2015
[3] Segnet: A deep convolutional encoder-decoder architecture for image segmentation, Badrinarayanan et al., IEEE PAMI 39(12) 2017

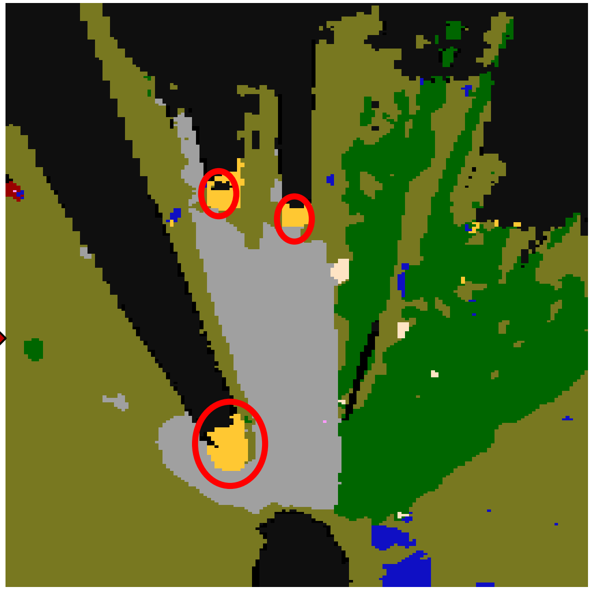
Semantic Grids: *Experimental Approach*

Frontal View *(RGB camera)*

Frontal View Ground-Truth
=> labelled by humans in datasets

Bird's Eye View Ground-Truth
⇒ *Frontal View GT **projected** using Point-Clouds (Bayesian Perception)*

⇒ *Densified by humans (different resolutions of images & point-clouds)*

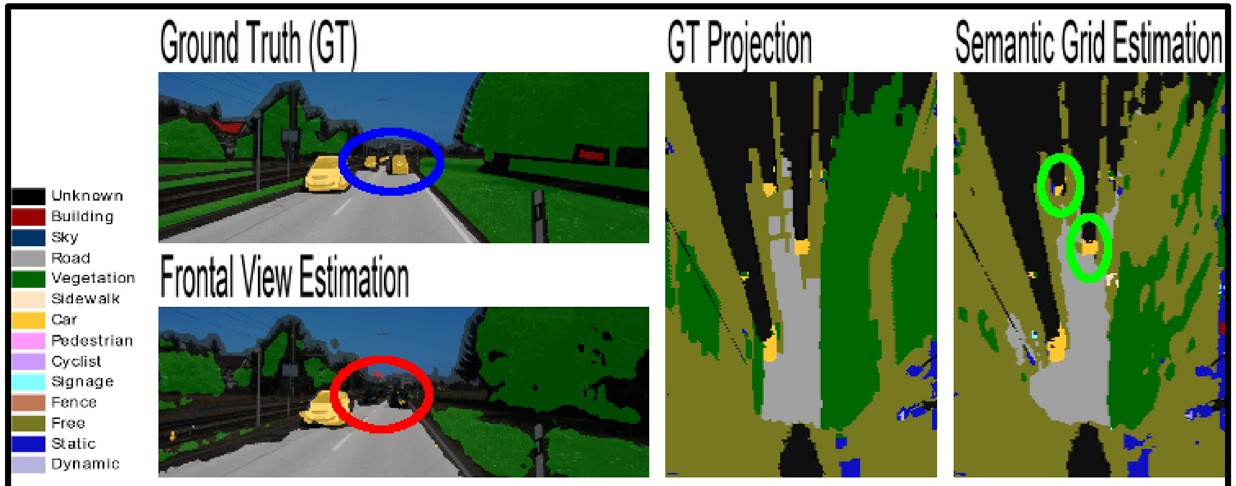


Semantic Grid Prediction
=> Dense structure obtained using hybrid integration

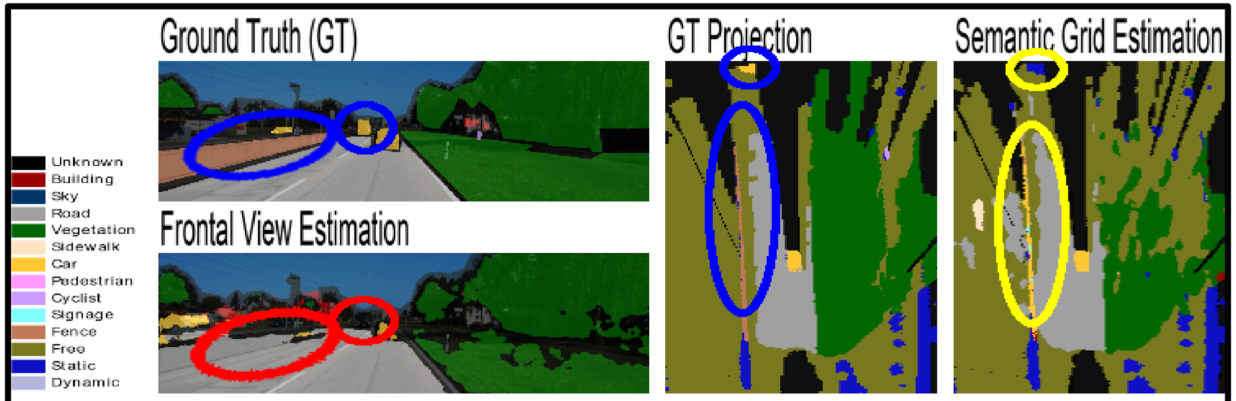
- Unknown
- Building
- Sky
- Road
- Vegetation
- Sidewalk
- Car
- Pedestrian
- Cyclist
- Signage
- Fence
- Free
- Static
- Dynamic

Labels

Semantic Grids: *Experimental Results & Current work*



Cars not detected in frontal view,
But they are recognized in semantic grid
with the help of dynamic occupancy grids.



Fence not detected in frontal view,
but detected as an obstacle class in Semantic Grid Estimation
with the help of dynamic occupancy grid.

The truck is not recognized in Frontal View,
but it is recognized as an obstacle in semantic view
with the help of dynamic occupancy grid.

Current Work

- Improve accuracy with more dense training datasets
- Implementation on embedded systems for real-time process
- Adaptation to bad weather conditions

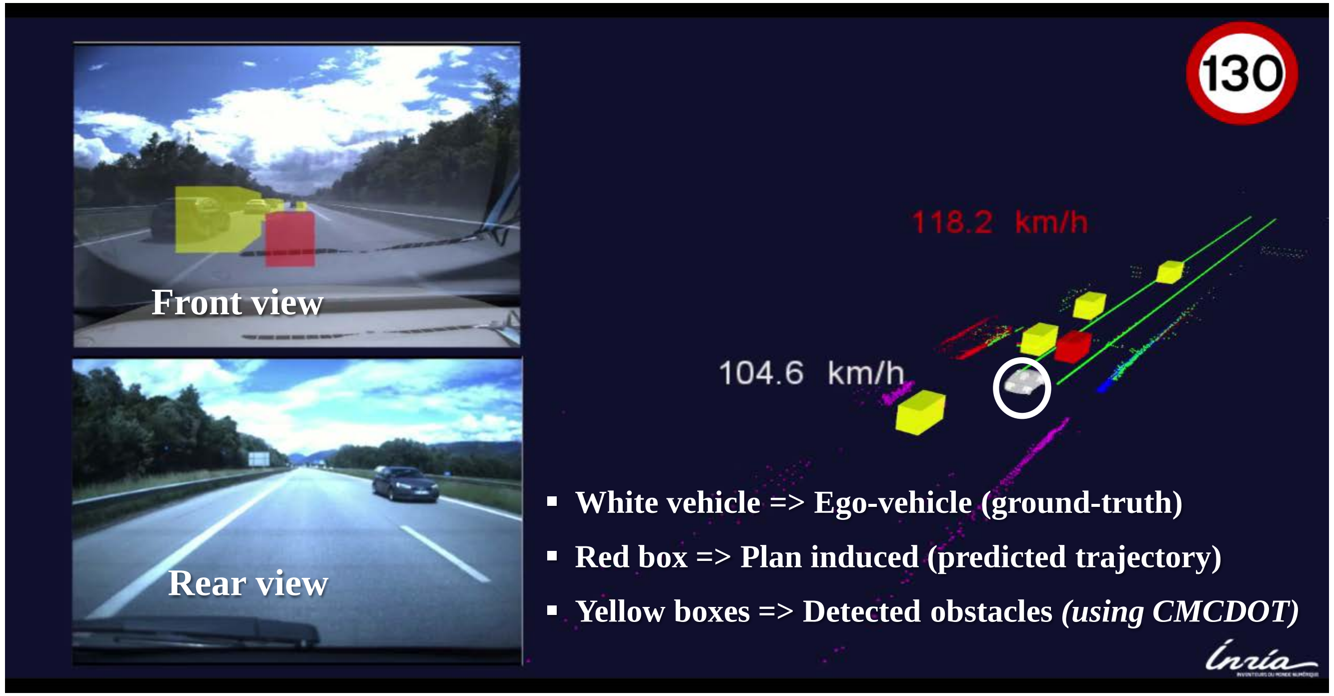
Car detected wrong in frontal view.
Corrected in Semantic Grid thanks to dynamic occupancy grid input.

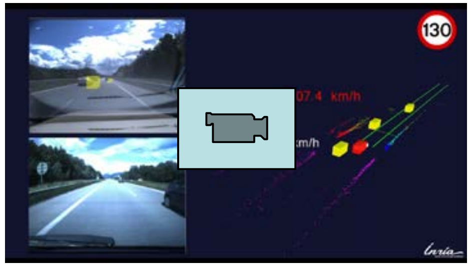
Paradigm 3: Models improvements using Machine Learning
Decision Level – Learning Driving Skills for AD

Step 1: Driver behavior modeling

[Sierra Gonzalez et al, ICRA 2018]

- **Learn the model parameters** automatically from driving demonstrations (real driving data) using **Inverse Reinforcement Learning**
- Driver behaviors are modelled using a **Cost function** $\mathcal{C}(s) = \sum_{i=1}^{K} w_i \cdot f_i(s)$ which is assumed linear on a **set of K hand-crafted features** (e.g. Lane index preferences, Deviation from desired velocity, Time-to-collision to frontal targets, Time-gap to rear targets)
- The obtained models can be leverage to **Predict human behaviors** and to **Generate human-like plans for the ego vehicle**
- A **training set** containing *"interesting highway vehicle interactions"* was constructed out of 20 minutes of highway driving data, and used to automatically learn the balance between features





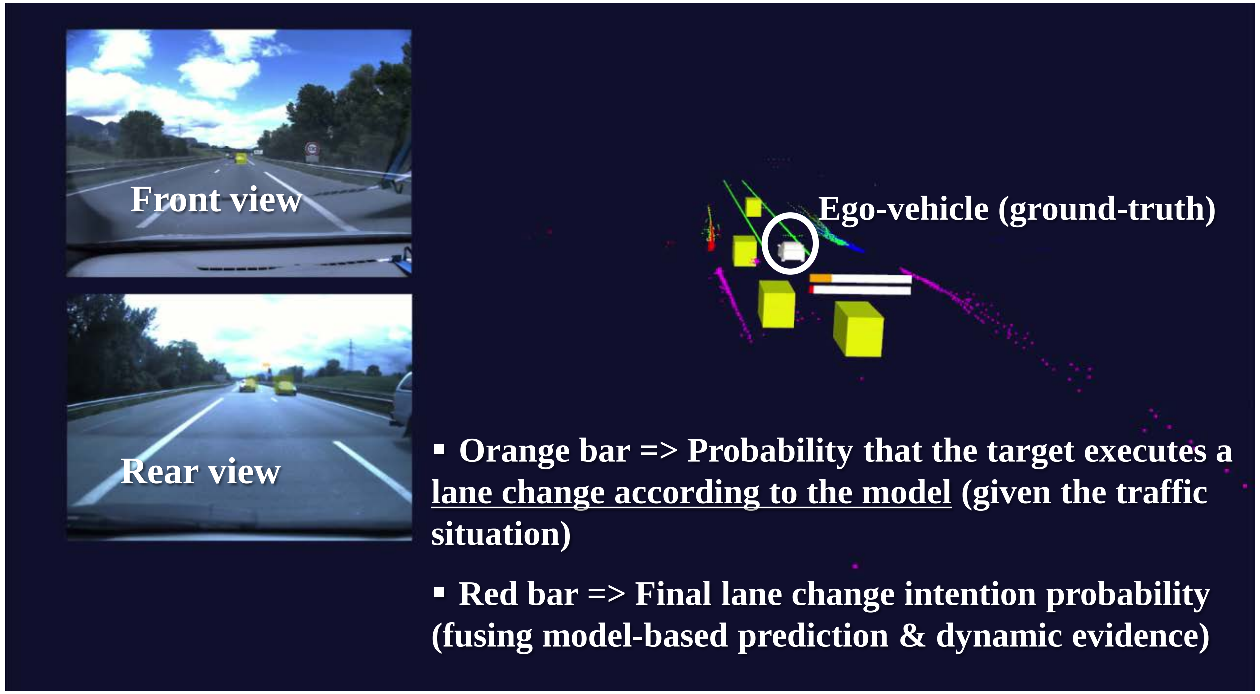
Comparison between demonstrated behavior in test set and behavior induced by the learned model

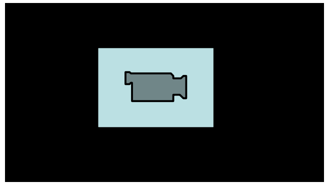
Paradigm 3: Models improvements using Machine Learning

Decision Level – Learning Driving Skills for AD

Step 2: Motion Prediction & Decision-making

- A **realistic Human-like Driver Model** can be exploited to **Predict the long-term evolution** (10s and beyond) of traffic scenes *[Sierra Gonzalez et al., ITSC 2016]*
- For the **short/mid-term**, both the **Dynamics of the target** and the **Driver model** provide useful information to **determine future behaviors**
- A **probabilistic model fuses *Model-based predictions* with *Dynamic evidence*** to produce robust **lane change intention estimations** in highway scenes *[Sierra Gonzalez et al., ICRA 2017]*





Comparison between demonstrated behavior in test set and behavior induced by the learned model

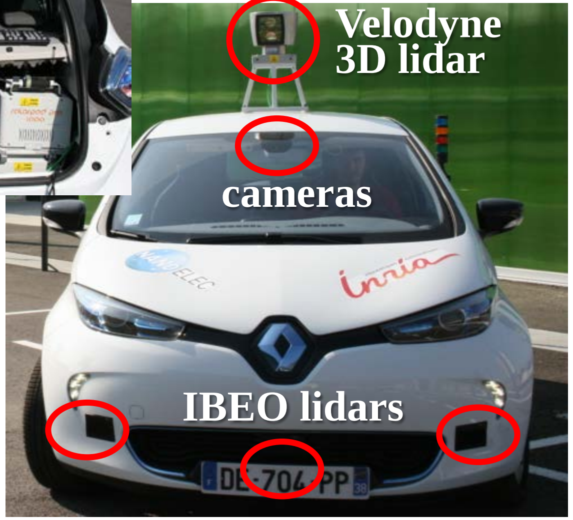
Experimental Vehicles & Connected Perception Units

Toyota Lexus

ROS

RT-Maps
under development

Renault Zoé



Connected Perception Unit (V2X communication)

Same embedded perception systems than in vehicles

=> Exchanging only relevant information (e.g. Risk parameters)

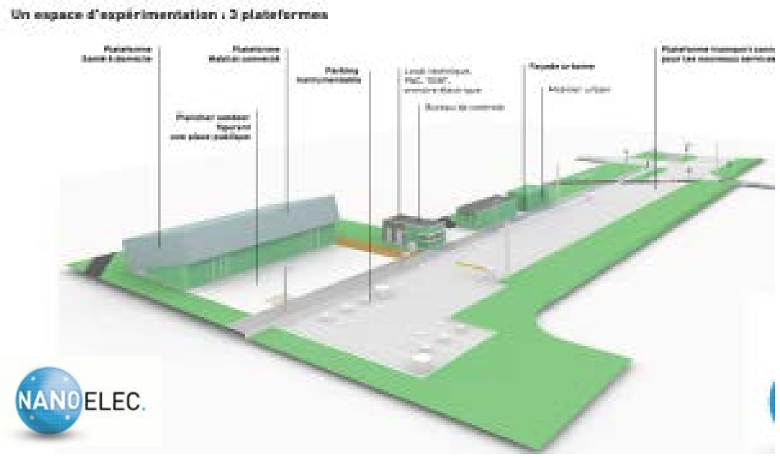
Nvidia GTX Titan X
Generation Maxwell

Nvidia GTX Jetson TK1
Generation Maxwell

Nvidia GTX Jetson TX1
Generation Maxwell

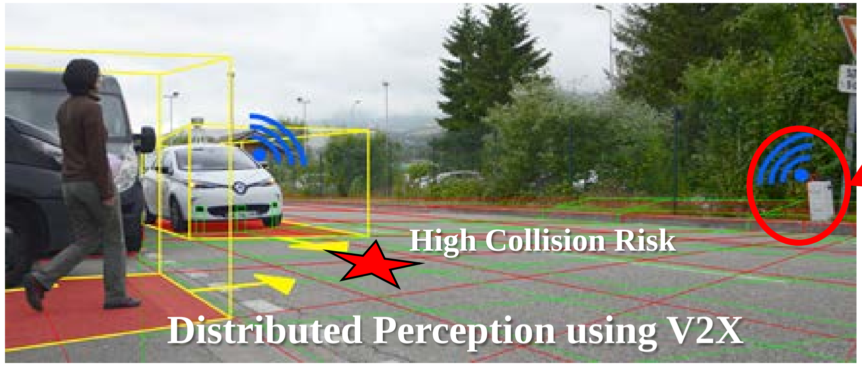
Experimental Areas

- **Protected experimental area** *=> Testing Autonomous Driving L3 & L4*



Crash test track

Connected
Perception Unit



- **Open real traffic (Urban & Highway)** *=> Testing Autonomous Driving L2 (ADAS)*

Summary & On going work

❑ Perception & Decision-making & Control integration (cooperation with industry)

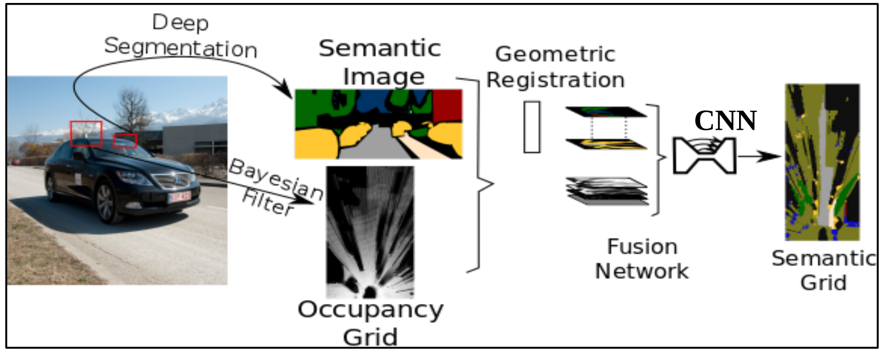
Autonomous Shuttles
(~15 km/h)

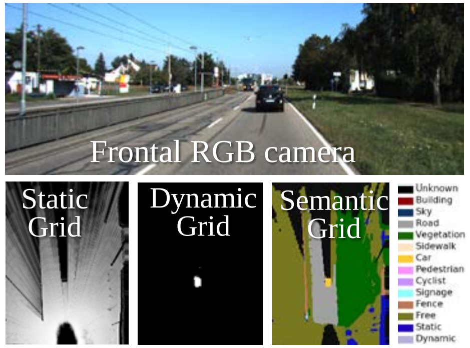
Autonomous Bus (Iveco)
(up to 70 km/h, Urban traffic)

Autonomous Renault Zoe
(up to 70 km/h, Urban traffic)

- Various Dynamics & Motion constraints & Contexts
- Adapted "Collision Risk" & "Collision avoidance maneuvers" (Risk & Maneuver characterization)
- **Cooperation IRT Nanoelec, Renault, Iveco …**

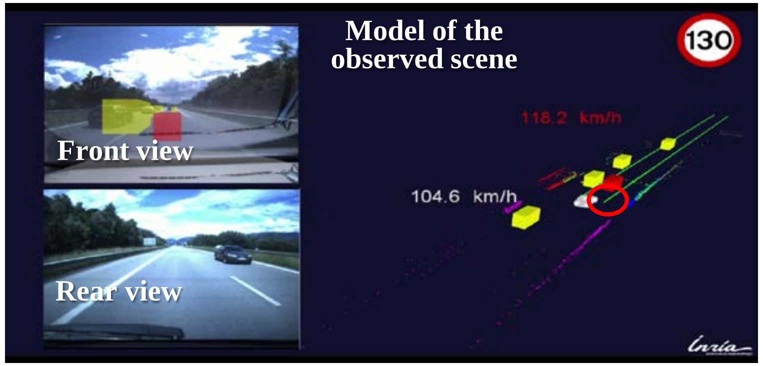
❑ Models enrichment using Semantic Segmentation & Deep Learning



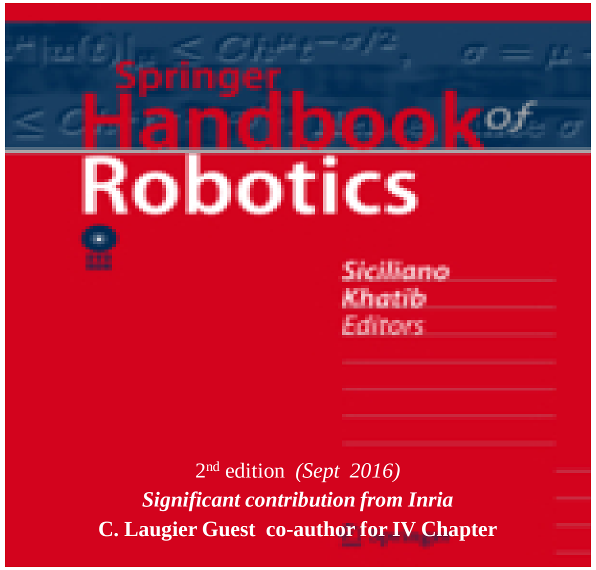


- "Semantic Grids" concept
- Improved scene understanding & decision-making
- **Cooperation Toyota**
- *1 Patent & 3 publications (IROS'18, ICARCV'18, Unmanned System journal 2019)*

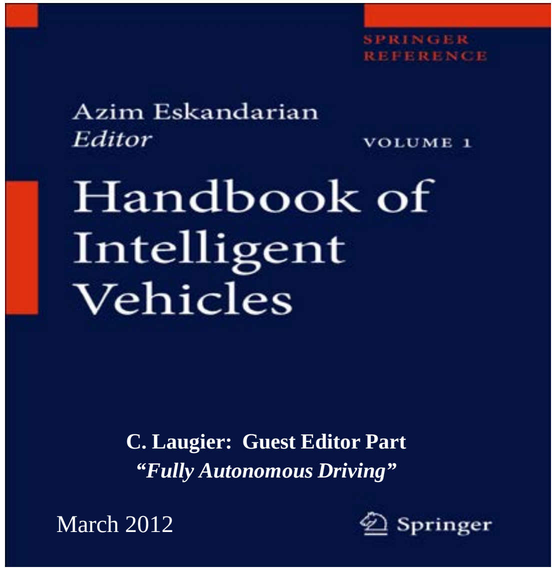
❑ Learning Driving Skills (Prediction & Planning) for Autonomous Driving



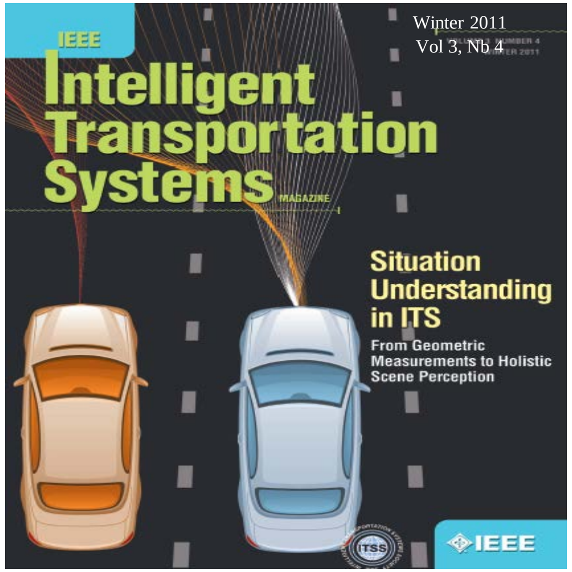
- **Driver Behavior modeling** using Driving dataset & Inverse Reinforcement Learning => ***Human-like Driver Model*** *(for mixed traffic)*
- **Motion Prediction & Driving Decision-making for AD** performed *by combining "learned Driver models" & "Dynamic evidences"*
- **Cooperation Toyota**
- *2 Patents & 3 publications (ITSC 2016, ICRA 2017, ICRA 2018) & PhD Thesis 2019*



2nd edition (Sept 2016)
Significant contribution from Inria
C. Laugier Guest co-author for IV Chapter



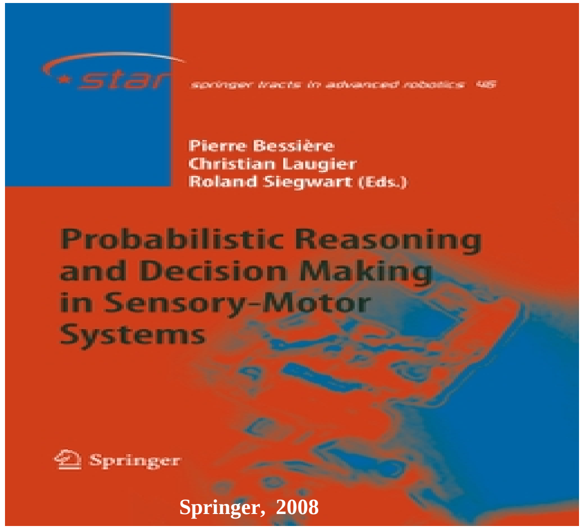
C. Laugier: Guest Editor Part
"Fully Autonomous Driving"
March 2012



Winter 2011
Vol 3, Nb 4

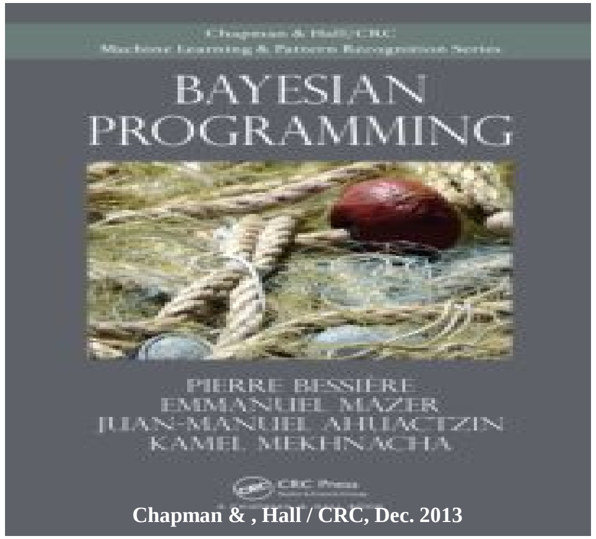
July 2013
Guest Editors:
C. Laugier & J. Machan

Thank You



Springer, 2008

IEEE RAS Technical Committee on "AGV & ITS"
Numerous Workshops & Special issues since 2002
=> Membership open !!



Chapman & , Hall / CRC, Dec. 2013